# Chapter 2

# A Review on Differentiation

Reading: Spivak pp. 15-34, or Rudin 211-220

## 2.1 Differentiation

Recall from 18.01 that

**Definition 2.1.1.** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is* **differentiable** *at $a \in \mathbb{R}^n$ if there exists a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 0 \tag{2.1}$$

The norm in Equation 2.1 is essential since $f(a+h) - f(a) - \lambda(h)$ is in $\mathbb{R}^m$ and $h$ is in $\mathbb{R}^n$.

**Theorem 2.1.2.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then there is a* **unique** *linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ that satisfies Equation* (2.1). *We denote $\lambda$ to be $Df(a)$ and call it the* **derivative** *of $f$ at $a$*

*Proof.* Let $\mu : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \mu(h)|}{|h|} = 0 \tag{2.2}$$

and $d(h) = f(a+h) - f(a)$, then

$$\lim_{h \to 0} \frac{|\lambda(h) - \mu(h)|}{|h|} = \lim_{h \to 0} \frac{|\lambda(h) - d(h) + d(h) - \mu(h)|}{|h|} \tag{2.3}$$

$$\leq \lim_{h \to 0} \frac{|\lambda(h) - d(h)|}{|h|} + \lim_{h \to 0} \frac{|d(h) - \mu(h)|}{|h|} \tag{2.4}$$

$$= 0. \tag{2.5}$$

Now let $h = tx$ where $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$, then as $t \to 0$, $tx \to 0$. Thus, for $x \neq 0$, we have

$$\lim_{t \to 0} \frac{|\lambda(tx) - \mu(tx)|}{|tx|} = \frac{|\lambda(x) - \mu(x)|}{|x|} \tag{2.6}$$

$$= 0 \tag{2.7}$$

Thus $\mu(x) = \lambda(x)$. □

Although we proved in Theorem 2.1.2 that if $Df(a)$ exists, then it is unique. However, we still have not discovered a way to find it. All we can do at this moment is just by guessing, which will be illustrated in Example 1.

**Example 1.** *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by*

$$g(x, y) = \ln x \tag{2.8}$$

**Proposition 2.1.3.** *$Dg(a, b) = \lambda$ where $\lambda$ satisfies*

$$\lambda(x, y) = \frac{1}{a} \cdot x \tag{2.9}$$

*Proof.*

$$\lim_{(h,k) \to 0} \frac{|g(a+h, b+k) - g(a, b) - \lambda(h, k)|}{|(h, k)|} = \lim_{(h,k) \to 0} \frac{|\ln(a+h) - \ln(a) - \frac{1}{a} \cdot h|}{|(h, k)|} \tag{2.10}$$

Since $\ln'(a) = \frac{1}{a}$, we have

$$\lim_{h \to 0} \frac{|\ln(a+h) - \ln(a) - \frac{1}{a} \cdot h|}{|h|} = 0 \tag{2.11}$$

Since $|(h,k)| \geq |h|$, we have

$$\lim_{(h,k) \to 0} \frac{|\ln(a+h) - \ln(a) - \frac{1}{a} \cdot h|}{|(h,k)|} = 0 \tag{2.12}$$

$\square$

**Definition 2.1.4.** *The* **Jacobian matrix** *of $f$ at $a$ is the $m \times n$ matrix of $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ with respect to the usual bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, and denoted $f'(a)$.*

**Example 2.** *Let $g$ be the same as in Example 1, then*

$$g'(a,b) = (\frac{1}{a}, 0) \tag{2.13}$$

**Definition 2.1.5.** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is* **differentiable on** *$A \subset \mathbb{R}^n$ if $f$ is diffrentiable at $a$ for all $a \in A$. On the other hand, if $f : A \to \mathbb{R}^m, A \subset \mathbb{R}^n$, then $f$ is called* **differentiable** *if $f$ can be extended to a differentiable function on some open set containing $A$.*

## 2.2 Properties of Derivatives

**Theorem 2.2.1.** *1. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a constant function, then $\forall a \in \mathbb{R}^n$,*

$$Df(a) = 0. \tag{2.14}$$

*2. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $\forall a \in \mathbb{R}^n$*

$$Df(a) = f. \tag{2.15}$$

*Proof.* The proofs are left to the readers □

**Theorem 2.2.2.** *If $g : \mathbb{R}^2 \to \mathbb{R}$ is defined by $g(x, y) = xy$, then*

$$Dg(a, b)(x, y) = bx + ay \tag{2.16}$$

*In other words, $g'(a, b) = (b, a)$*

*Proof.* Substitute $p$ and $Dp$ into L.H.S. of Equation 2.1, we have

$$\begin{aligned}
\lim_{(h,k)\to 0} \frac{|g(a+h, b+k) - g(a, b) - Dg(a, b)(h, k)|}{|(h, k)|} &= \lim_{(h,k)\to 0} \frac{|hk|}{|(h, k)|} && (2.17)\\
&\le \lim_{(h,k)\to 0} \frac{\max(|h|^2, |k|^2)}{\sqrt{h^2+k^2}} && (2.18)\\
&\le \sqrt{h^2+k^2} && (2.19)\\
&= 0 && (2.20)
\end{aligned}$$

□

**Theorem 2.2.3.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a$, and $g : \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $f(a)$, then the composition $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at $a$, and*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a) \tag{2.21}$$

*Proof.* Put $b = f(a)$, $\lambda = f'(a)$, $\mu = g'(b)$, and

$$u(h) = f(a+h) - f(a) - \lambda(h) \tag{2.22}$$

$$v(k) = g(b+k) - g(b) - \mu(k) \tag{2.23}$$

for all $h \in \mathbb{R}^n$ and $k \in \mathbb{R}^m$. Then we have

$$|u(h)| = \epsilon(h)|h| \tag{2.24}$$

$$|v(k)| = \eta(k)|k| \tag{2.25}$$

where

$$\lim_{h\to 0} \epsilon(h) = 0 \tag{2.26}$$

$$\lim_{k\to 0} \eta(k) = 0 \tag{2.27}$$

Given $h$, we can put $k$ such that $k = f(a+h) - f(a)$. Then we have

$$|k| = |\lambda(h) + u(h)| \leq [\|\lambda\| + \epsilon(h)]|h| \tag{2.28}$$

Thus,

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) - \mu(\lambda(h)) &= g(b+k) - g(b) - \mu(\lambda(h)) && (2.29) \\ &= \mu(k - \lambda(h)) + v(k) && (2.30) \\ &= \mu(u(h)) + v(k) && (2.31) \end{aligned}$$

Thus

$$\frac{|g \circ f(a+h) - g \circ f(a) - \mu(\lambda(h))|}{|h|} \leq \|\mu\|\epsilon(h) + [\|\lambda\| + \epsilon(h)]\eta(h) \tag{2.32}$$

which equals 0 according to Equation 2.26 and 2.27. □

**Exercise 1.** *(Spivak 2-8) Let $f : \mathbb{R} \to \mathbb{R}^2$. Prove that $f$ is differentiable at $a \in \mathbb{R}$ if and only if $f^1$ and $f^2$ are, and that in this case*

$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \end{pmatrix} \tag{2.33}$$

**Corollary 2.2.4.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$, then $f$ is differentiable at $a \in \mathbb{R}^n$ if and*

*only if each $f^i$ is, and*

$$\lambda'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \\ . \\ . \\ . \\ (f^m)'(a) \end{pmatrix}. \tag{2.34}$$

*Thus, $f'(a)$ is the $m \times n$ matrix whose $i$th row is $(f^i)'(a)$*

**Corollary 2.2.5.** *If $f, g : \mathbb{R}^n \to \mathbb{R}$ are differentiable at $a$, then*

$$D(f+g)(a) = Df(a) + Dg(a) \tag{2.35}$$

$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a) \tag{2.36}$$

$$D(f/g)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{[g(a)]^2}, \ g(a) \neq 0 \tag{2.37}$$

*Proof.* The proofs are left to the readers. □

## 2.3 Partial Derivatives

**Definition 2.3.1.** *If $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$, then the limit*

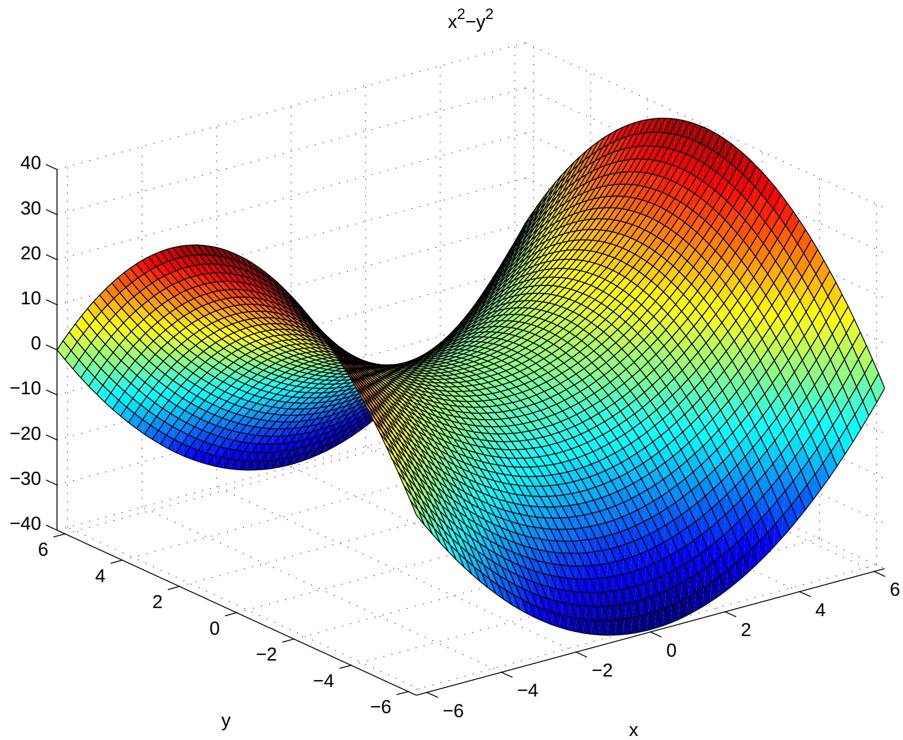
$$D_i f(a) = \lim_{h \to 0} \frac{f(a^1, ..., a^i + h, ..., a^n) - f(a^1, ..., a^n)}{h} \tag{2.38}$$

*is called the $i$th* **partial derivative** *of $f$ at $a$ if the limit exists.*

If we denote $D_j(D_i f)(x)$ to be $D_{i,j}(x)$, then we have the following theorem which is stated without proof. (The proof can be found in Problem 3-28 of Spivak)

**Theorem 2.3.2.** *If $D_{i,j}f$ and $D_{j,i}f$ are continuous in an open set containing $a$, then*

$$D_{i,j}f(a) = D_{j,i}f(a) \tag{2.39}$$




Partial derivatives are useful in finding the extrema of functions.

**Theorem 2.3.3.** *Let $A \subset \mathbb{R}^n$. If the maximum (or minimum) of $f : A \to \mathbb{R}$ occurs at a point $a$ in the interior of $A$ and $D_i f(a)$ exists, then $D_i f(a) = 0$.*

*Proof.* The proof is left to the readers. □

However the converse of Theorem 2.3.3 may not be true in all cases. (Consider $f(x, y) = x^2 - y^2$).

## 2.4 Derivatives

**Theorem 2.4.1.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a$, then $D_j f^i(a)$ exists for $1 \leq i \leq m, 1 \leq j \leq n$ and $f'(a)$ is the $m \times n$ matrix $(D_j f^i(a))$.*